HUMAN OXIDOREDUCTASE PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human oxidoreductase proteins and to the use of these sequences in the diagnosis, treatment, and prevention of neurological, autoimmune/inflammatory, reproductive, cell proliferative, endocrine, vesicle trafficking, and smooth muscle disorders.

BACKGROUND OF THE INVENTION

Many pathways of biogenesis and biodegradation require oxidoreductase (dehydrogenase or reductase) activity, coupled to reduction or oxidation of a cofactor, such as $NAD(P)^+/NAD(P)H$. (Newsholme, E.A. and Leech, A.R. (1983) Biochemistry for the Medical Sciences, John Wiley and Sons, Chichester, U.K. pp. 779-793.) NADH and NADPH bind canonical conserved amino acid sequence motifs. (Scrutton, N.S. et al. (1990) Nature, 343:38-43.) Reductase activity catalyzes the transfer of electrons between substrate(s) and cofactor(s) with concurrent oxidation of NADH and NADPH. The reverse dehydrogenase reaction catalyzes the reduction of NAD^+ and $NADP^+$. Electron transport is a general cellular process in which electrons generated from the oxidation of molecules such as NADH and $FADH_2$ are transferred, through the action of various enzymes, to a series of electron carriers. These electron carriers may act as electron donors for various reductive reactions in the cell or may transport their electrons to other electron carriers along an electron transport chain. The change in oxidation potential as electrons are passed along this chain generates energy which may be used by the cell.

The mitochondrial electron transport, or respiratory, chain is comprised of a series of enzyme complexes in the mitochondrial membrane. This transport chain is responsible for the transfer of electrons from NADH through a series of redox centers (electron carriers) within these complexes to oxygen, as well as for the coupling of this oxidation to the synthesis of ATP (oxidative phosphorylation). ATP provides the primary source of energy for driving a cell's many energy-requiring reactions. The key enzyme complexes in the respiratory chain are NADH-dependent ubiquinone oxidoreductase, succinate:ubiquinone oxidoreductase, cytochrome c_1-b oxidoreductase, cytochrome c oxidase, and ATP synthase. These complexes are located on the inner matrix side of the mitochondrial membrane with the exception of succinate:ubiquinone oxidoreductase, which is located on the cytosolic side. NADH-dependant ubiquinone oxidoreductase initiates the first step in the respiratory chain by accepting electrons from NADH

and passing them through a flavin molecule and several iron-sulfur centers to ubiquinone.

NADPH-dependent quinone oxidoreductases modulate mitochondrial electron transfer from flavoproteins to cytochromes. During the electron transfer reaction, ubiquinone-10 is reduced in the presence of NADPH to ubiquinol-10. Ubiquinone-10 is regenerated by cytochrome c_1-b reductase from ubiquinol-10 by electron transfer to cytochrome c. These reactions enable electron transfer from a lipid to an aqueous phase and the concomitant pumping of protons across the inner mitochondrial membrane.

Defects and altered expression of NADH-dependent ubiquinone oxidoreductase are associated with a variety of human diseases, including neurodegenerative diseases, myopathies, and cancer. (Singer, T.P. et al. (1995) Biochim. Biophys. Acta 1271:211-219; and Selvanayagam, P. and Rajaraman, S. (1996) Lab. Invest. 74:592-599.) In addition, NADH-dependent ubiquinone oxidoreductase reduction of the quinone moiety in chemotherapeutic agents such as doxorubicin is believed to contribute to the antitumor activity and/or mutagenicity of these drugs. (Akman, S.A. et al. (1992) Biochemistry 31:3500-3506.)

Many enzymes involved in biogenesis and biodegradation of neurotransmitters and other signaling molecules have been isolated and identified as being members of the short-chain NAD^+/NADH-dependent alcohol dehydrogenase family. Inactivation of neurotransmitters is absolutely essential for the correct functioning of the nervous system, as the use of experimental neuroactive drugs has shown. (Avery, L. and Horvitz, H.R. (1990) J. Ex. Zool. 253:263-270.) Epigenetic or genetic defects in neurotransmitter metabolic pathways can result in a spectrum of disease states in different tissues including Parkinson disease and inherited myoclonus. (McCance, K.L. and Huether, S.E. (1994) Pathophysiology, Mosby-Year Book, Inc., St. Louis, MO pp. 402-404; and Gundlach, A.L. (1990) FASEB J. 4:2761-2766.)

Other examples of enzymes with intracellular oxidoreductase activity include asparagine amidohydrolase, biotin-amide amidohydrolase, and quinone oxidoreductases. N-terminal asparagine amidohydrolase catalyzes the deamidation of the N-terminal asparagine residue of a protein to aspartate. This is the first step in the protein destabilizing pathway of protein degradation according to the N-end rule which sates that the in vivo half-life of a protein is related to the identity of its N-terminal residue. The N-end rule specifies a set of N-termini which contain different destabilizing residues. (Grigoryev, S., et al. (1996) J. Biol. Chem. 271:28521-28532.) Aspartate is considered a secondary N-terminal destabilizing residue. N-terminal aspartate is then enzymatically converted to arginine, a primary destabilizing residue.

Biotin is the coenzyme for four carboxylase enzyme reactions necessary for metabolism. Biotin is the intermediary for the transfer of one molecule of carbon dioxide to the substrate. Biotin-dependent enzymes include acetyl CoA carboxylase, which is the first step in lipid

biosynthesis, and pyruvate carboxylase which incorporates carbon into pyruvate to form oxaloacetate. Biocytin (biotin-ξ-lysine) is a product of the carboxylase activity. Biotin-amide amidohydrolase (also termed biotinidase, EC 3.5.1.12) catalyzes the conversion of biocytin to biotin and lysine, thereby recycling the vitamin. Biotinidase deficiency has been linked to the late-onset form of multiple carboxylase deficiency, which includes symptoms of seizure, hypotonia, optic atrophy, alopecia, ketolactic acidosis, and organic acidurea.

Cytochromes are electron-transferring proteins that contain a heme prosthetic group, a porphyrin ring containing a tightly bound iron atom. In addition to their role in respiration, cytochromes also act as oxidoreductases in such diverse cellular processes as photosynthesis, fatty acid metabolism, and neurotransmitter biosynthesis. The heme iron atom serves as the actual electron carrier by changing from the ferric to the ferrous oxidation state when accepting an electron. Cytochromes accept electrons from one substrate such as NADH or ascorbate and donate them to other electron carriers such as other cytochromes, ubiquinone, or semidehydroascorbic acid. (Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York, NY, pp. 759-770, 786-797; Sperling, P. et al. (1995) Eur. J. Biochem. 232:798-805; and Online Mendelian Inheritance in Man (OMIM) *600019 Cytochrome b561, CYB561.)

Cytochrome b5 is an electron donor in membrane-linked redox enzyme systems involved in lipid and drug metabolism. Cytochrome b5 has been found in Golgi, plasma, outer mitochondrial, endoplasmic reticulum (ER), and microbody membranes. Conserved amino acids in cytochrome b5 include eight invariant amino acids at W34, H51, P52, G53, G54, G63, F70, and H74, and fifteen conserved amino acids at L24, I35, S36, V41, Y42, N43, T45, W47, A48, L58, D65, T67, L85, T87, and G88 (numbering based on the sunflower cytochrome b5/delta-6 desaturase fusion protein; GI 1040729, Sperling, supra). The invariant residues H51PGG are involved in heme-binding. Cytochrome b5-like domains have also been found linked to other enzymes. For example, cytochrome b5-like domains are part of delta-9 fatty acid desaturases in yeast and Histoplasma capsulatum, nitrate reductase, sulfite reductase, flavocytochrome b2, Arabidopsis thaliana acyl lipid desaturase, and Borago officinalis (borage) and Helianthus annuus (sunflower) delta-6 desaturases. (Sperling, supra; Sayanova, O. et al (1997) Proc. Natl. Acad. Sci. USA 94:4211-4216; and Mitchell, A.G. and Martin, C.E. (1997) J. Biol. Chem. 272:28281-28288.)

Delta-6 desaturases are enzymes that catalyze the elongation and desaturation of fatty acids, for example, the conversion of the essential fatty acid linoleic acid to γ-linolenic acid. The delta-6 desaturated fatty acids have roles in maintenance of membrane structure and function, in the regulation of cholesterol synthesis and transport, in the prevention of water loss from the skin, and as precursors of eicosanoids, including prostaglandins and leucotrienes (Sayanova, supra).

The desaturation reactions are often carried out by a complex consisting of a desaturase, a cytochrome b5, and a cytochrome b5-reductase, on the cytosolic face of the ER membrane. (Leikin, A. and Shinitzky, M. (1994) Biochim. Biophys. Acta 1211:150-155.) Marzo et al. (1996; Biochim. Biophys. Acta 1301:263-272) suggest that two delta-6 desaturase activities may exist in humans and that each may be specific to different fatty acid substrates. The borage and sunflower cytochrome b5/desaturase fusion proteins contain an N-terminal cytochrome b5 domain which includes a putative heme binding site and other conserved amino acids followed by a delta-6 desaturase domain containing three conserved histidine boxes of general sequence $HX_{2(3)}[XH]H$ (where X=any amino acid). The histidines in the histidine boxes may act as metal-chelating ligands and contribute to the binding of oxygen in the reaction center. The borage cytochrome b5/desaturase fusion protein may be localized to the ER. Fusion of the cytochrome b5 domain to the desaturase may facilitate more efficient electron transfer. (Sayanova, supra; Sperling, supra.)

Delta-6 desaturases and their products are associated with several disorders. Nakada et al. (1990; Neuroreport 1:153-155) analyzed the membrane fatty acid composition of Alzheimer's diseased brain tissue and found evidence of abnormalities in delta-6 desaturation. Decreased delta-6 desaturase activity is observed in aging, stress, diabetes, eczema, and some infections. Increased catabolism of γ-linolenic acid (the product of delta-6 desaturase) results from more rapid cell division due, for example, to cancer or inflammation. Both decreased delta-6 desaturase activity and increased γ-linolenic acid catabolism can lead to a deficiency of γ-linolenic acid. γ-linolenic acid can be effective in treating atopic eczema, mastalgia, diabetic neuropathy, viral infections, and some types of cancer (Sayanova, supra).

Cytochrome b561 acts as an electron channel in the biosynthesis of several catecholamine and peptide neurotransmitters and is found in catecholamine and neuropeptide secretory vesicles of the adrenal medulla, pituitary gland, and other neuroendocrine tissues. Cytochrome b561 mRNA has been detected in brain, placenta, lung, pancreas, and kidney. Cytochrome b561 supplies reducing equivalents to two monooxygenases: dopamine beta-hydroxylase in chromaffin granules and peptidyl-glycine alpha-amidating monooxygenase in neurosecretory vesicles. Cytochrome b561 catalyzes the transfer of electrons from the cytoplasmic donor ascorbate across the phospholipid bilayer to the luminal acceptor semidehydroascorbic acid. The continuously regenerated ascorbate donates electrons to the monooxygenases. (OMIM, supra; and Srivastava, M. et al. (1994) Biochem. J. 303:915-921.) The human cytochrome b561 has been described as having five transmembrane spans, and the bovine form six transmembrane spans. Six residues in cytochrome b561, which are conserved between the human and bovine forms, are proposed to play a role in heme binding. These residues are H87, W136, G139, F140, H160, and F163 (numbering based on the human cytochrome b561). (GI 1345640; Srivastava, (1994) supra.)

Cytochrome b561 enzymes have been found in many organisms including human, bovine, Xenopus laevis, and mouse. The Xenopus laevis and mouse forms have been reported to be developmentally regulated. (Srivastava, M. (1995) J. Biol. Chem. 270:22714-22720.)

The fact that expression of cytochrome b561 is restricted to neuroendocrine tissues suggested to Srivastava et al. (1994; supra) that cytochrome b561 may have a role in neuroendocrine disorders. Srivastava (1995, supra) suggested possible roles for cytochrome b561 in development, cell differentiation, and carcinogenesis. No expression of cytochrome b561 mRNA was observed in the B-cell lymphoma cells Burkitt's and Daudi. Cytochrome b561 expression was increased in colon cancer cell lines, T cell lymphomas, and undifferentiated cell lines when compared to expression in peripheral blood leukocytes (Srivastava (1995) supra).

The discovery of new human oxidoreductase proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, treatment, and prevention of neurological, autoimmune/inflammatory, reproductive, cell proliferative, endocrine, vesicle trafficking, and smooth muscle disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, human oxidoreductase proteins, referred to collectively as "HORP" and individually as "HORP-1", "HORP-2", HORP-3", HORP-4", HORP-5", and "HORP-6". In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 (SEQ ID NO:1-6), and fragments thereof.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-6 and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof, as well as an isolated and purified polynucleotide having a sequence which is

complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12 (SEQ ID NO:7-12) and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:7-12 and fragments thereof as well as an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:7-12 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide encoding the polypeptide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof, as well as a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of HORP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of HORP, the method comprising administering to a subject in

need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof.

The invention also provides a method for detecting a polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof in a biological sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of the polynucleotide sequence encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-6 and fragments thereof to at least one of the nucleic acids of the biological sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide encoding the polypeptide in the biological sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

BRIEF DESCRIPTION OF THE FIGURES AND TABLES

Figures 1A, 1B, and 1C show the amino acid sequence alignment between HORP-5 (008879; SEQ ID NO:5) and sunflower cytochrome b5/desaturase fusion protein (GI 1040729; SEQ ID NO: 13).

Figures 2A and 2B show the amino acid sequence alignment between HORP-6 (2274011; SEQ ID NO:6) and human cytochrome b561 (GI 1345640; SEQ ID NO:14). The alignments were produced using the multisequence alignment program of LASERGENE™ software (DNASTAR Inc, Madison WI).

The first column of Table 1 shows the polypeptide sequence identifiers, SEQ ID NO:1-6. The second column shows the nucleotide sequence identifiers, SEQ ID NO:7-12, of the consensus sequences which encode HORP. The third column lists the Incyte Clone number in which nucleic acids encoding each HORP were first identified. The fourth column lists the tissue library from which the clone was isolated. The fifth column lists the overlapping and/or extended nucleic acid sequences which were used to derive the consensus sequences SEQ ID NO:7-12.

The first column of Table 2 lists the polypeptide sequence identifiers. The second column shows the number of residues of SEQ ID NO:1-6. The third column lists the potential phosphorylation sites in SEQ ID NO:1-6, and the fourth column lists potential N-glycosylation sites. The fifth column lists any significant protein family signature or ligand/substrate binding motif present in SEQ ID NO:1-6. The sixth column indicates the identity of the protein. The seventh column describes the method of analysis or algorithm(s) used to identify the polypeptide.

The first column of Table 3 lists the polynucleotide sequence identifiers (SEQ ID NO:7-

12). The second column lists the tissue expression of HORP and fraction of total tissue which express SEQ ID NO:7-12. The third column lists the disease class and fraction of total disease tissues that express SEQ ID NO:7-12. The fourth column lists the vector used to subclone the cDNA library.

The first column of Table 4 list the nucleotide sequence identifiers (SEQ ID NO:7-10) for HORP-1, HORP-2, HORP-3, and HORP-4. The second column lists the corresponding Incyte Clone number. The third column lists the description of the tissue library from which the clone was derived.

Table 5 summarizes the programs, algorithms, databases, and qualifying scores used to analyze HORP. The first column of Table 5 shows the tool, program, or algorithm; the second column, the database, the third column, a brief description, and the fourth column (where applicable), scores for determining the strength of a match between two sequences (the higher the value, the greater the homology).

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular methodology, protocols, cell lines, vectors, and reagents described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods, devices, and materials are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, vectors, and methodologies which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"HORP," as used herein, refers to the amino acid sequences of substantially purified HORP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist," as used herein, refers to a molecule which, when bound to HORP, increases or prolongs the duration of the effect of HORP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of HORP.

An "allelic variant," as this term is used herein, is an alternative form of the gene encoding HORP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding HORP, as described herein, include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as HORP or a polypeptide with at least one functional characteristic of HORP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HORP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HORP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HORP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HORP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

The terms "amino acid" or "amino acid sequence," as used herein, refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of HORP which are preferably at least 5

to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of HORP. Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification," as used herein, relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art. (See, e.g., Dieffenbach, C.W. and G.S. Dveksler (1995) PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, NY, pp.1-5.)

The term "antagonist," as it is used herein, refers to a molecule which, when bound to HORP, decreases the amount or the duration of the effect of the biological or immunological activity of HORP. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of HORP.

As used herein, the term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind HORP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant," as used herein, refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense," as used herein, refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

As used herein, the term "biologically active," refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HORP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity," as used herein, refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" binds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition comprising a given amino acid sequence," as these terms are used herein, refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HORP or fragments of HORP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts, e.g., NaCl, detergents, e.g.,sodium dodecyl sulfate (SDS), and other components, e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.

"Consensus sequence," as used herein, refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using XL-PCR™ (Perkin Elmer, Norwalk, CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW™ Fragment Assembly system (GCG, Madison, WI). Some sequences have been both extended and assembled to produce the consensus sequence.

As used herein, the term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding HORP, by Northern analysis is indicative of the presence of nucleic acids encoding HORP in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding HORP.

A "deletion," as the term is used herein, refers to a change in the amino acid or nucleotide

sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative," as used herein, refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

The term "similarity," as used herein, refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" or "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MegAlign™ program (DNASTAR, Inc., Madison WI). The MegAlign™ program can create alignments between two or more sequences according to different methods, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity

between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs), as described herein, are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355.)

The term "humanized antibody," as used herein, refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization," as the term is used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

As used herein, the term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" or "addition," as used herein, refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray," as used herein, refers to an arrangement of distinct polynucleotides arrayed on a substrate, e.g., paper, nylon or any other type of membrane, filter, chip, glass slide, or any other suitable solid support.

The terms "element" or "array element" as used herein in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate," as it appears herein, refers to a change in the activity of HORP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HORP.

The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a

nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which, comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:7-12, for example, as distinct from any other sequence in the same genome. For example, a fragment of SEQ ID NO:7-12 is useful in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:7-12 from related polynucleotide sequences. A fragment of SEQ ID NO:7-12 is at least about 15-20 nucleotides in length. The precise length of the fragment of SEQ ID NO:7-12 and the region of SEQ ID NO:7-12 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment. In some cases, a fragment, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" or "operably linked," as used herein, refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide," as used herein, refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. As used herein, the term "oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA), as used herein, refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell. (See, e.g., Nielsen, P.E. et al. (1993) Anticancer Drug Des. 8:53-63.)

The term "sample," as used herein, is used in its broadest sense. A biological sample suspected of containing nucleic acids encoding HORP, or fragments thereof, or HORP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from

a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a solid support; a tissue; a tissue print; etc.

As used herein, the terms "specific binding" or "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

As used herein, the term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified," as used herein, refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution," as used herein, refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Transformation," as defined herein, describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of HORP polypeptides, as used herein, refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g.,

replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE™ software.

The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to HORP. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

THE INVENTION

The invention is based on the discovery of new human oxidoreductase proteins, HORP, the polynucleotides encoding HORP, and the use of these compositions for the diagnosis, treatment, or prevention of neurological, autoimmune/inflammatory, reproductive, cell proliferative, endocrine, vesicle trafficking, and smooth muscle disorders. Table 1 summarizes the sequence identification numbers, identifying clone numbers, and libraries of HORP .

As shown in Table 2, each HORP has been characterized with regard to its chemical and structural similarity with oxidoreductases. In Table 3, northern analysis shows the expression of nucleic acid sequences which encode HORP in various libraries, at least 42% of which are immortalized or cancerous, and at least 16% of which involve immune response. Of particular note is the expression of HORP-1, HORP-2, HORP-3, and HORP-4 in reproductive, gastrointestinal, cardiovascular, and nervous tissues.

As shown in Figures 1A, 1B, and 1C, HORP-5 has chemical and structural similarity with sunflower cytochrome b5/desaturase fusion protein (GI 1040729; SEQ ID NO:13). In particular, HORP-5 and sunflower cytochrome b5/desaturase fusion protein share 23% identity. HORP-5 and sunflower cytochrome b5/desaturase fusion protein share a potential heme-binding site, located at H53PGG in HORP-5; seven cytochrome b5 superfamily invariant residues, located at

H53, P54, G55, G56, G65, F72, and H76 in HORP-5; and nine cytochrome b5 superfamily conserved residues, located at V43, Y44, N45, T47, W49, D67, T69, L89, and G92 in HORP-5. HORP-5 and sunflower cytochrome b5/desaturase fusion protein have similar amino acid residues at four cytochrome b5 superfamily conserved sites, located at I26, L37, S50, and I60 in HORP-5. HORP-5 and sunflower cytochrome b5/desaturase fusion protein share three potential desaturase histidine boxes; located at H180DYGH, H217FQHH, and Q382IEHH in HORP-5. Northern analysis shows the expression of SEQ ID NO:11 in various libraries, at least 59% of which are immortalized or cancerous, at least 26% of which involve immune response, and at least 23% of which are fetal or proliferating cell or tissues. Of particular note is the expression of SEQ ID NO:11 in male and female reproductive, nervous, cardiovascular, and endocrine tissues. A fragment of SEQ ID NO:11 from about nucleotide 216 to about nucleotide 233 is useful, for example, as a hybridization probe.

As shown in Figures 2A and 2B, HORP-6 has chemical and structural similarity with human cytochrome b561 (GI 1345640; SEQ ID NO:14). In particular, HORP-6 and human cytochrome b561 share 34% identity. Both HORP-6 and human cytochrome b561 contain five potential transmembrane regions. HORP-6 and human cytochrome b561 share four potential cytochrome b561 heme-binding residues, located at H86, G138, F139, and H159 in HORP-6. Northern analysis shows the expression of SEQ ID NO:12 in various libraries, at least 48% of which are immortalized or cancerous, at least 32% of which involve immune response, and at least 12% are fetal or proliferating cells or tissues. Of particular note is the expression of SEQ ID NO:12 in male and female reproductive, gastrointestinal, cardiovascular, nervous, and endocrine tissues. A fragment of SEQ ID NO:12 from about nucleotide 332 to about nucleotide 349 is useful, for example, as a hybridization probe.

The invention also encompasses HORP variants. A preferred HORP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the HORP amino acid sequence, and which contains at least one functional or structural characteristic of HORP.

The invention also encompasses polynucleotides which encode HORP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:7-12, which encodes HORP.

The invention also encompasses a variant of a polynucleotide sequence encoding HORP. In particular, such a variant polynucleotide sequence will have at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HORP. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural

characteristic of HORP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HORP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HORP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HORP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring HORP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HORP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HORP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HORP and HORP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HORP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:7-12 or fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will

ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE® (US Biochemical Corp., Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham, Chicago, IL), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE™ Amplification System (GIBCO-BRL, Gaithersburg, MD). Preferably, sequence preparation is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, NV), Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA) and the ABI Catalyst 800 (Perkin Elmer). Sequencing is then carried out using either ABI 373 or 377 DNA Sequencing Systems (Perkin Elmer) or capillary electrophoresis (Molecular Dynamics). The resulting sequences are analyzed using a variety of alogorithms which are well known in the art. (See, e.g., Ausubel, supra, ch. 7.7; and Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley

VCH, Inc., New York, NY, pp. 856-853.)

The nucleic acid sequences encoding HORP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-306). Additionally, one may use PCR, nested primers, and PromoterFinder™ libraries to walk genomic DNA (Clontech, Palo Alto, CA). This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO™ 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., Genotyper™ and Sequence Navigator™, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA

fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HORP may be cloned in recombinant DNA molecules that direct expression of HORP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HORP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HORP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding HORP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, and Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232.) Alternatively, HORP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A Peptide Synthesizer (Perkin Elmer). Additionally, the amino acid sequence of HORP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g, Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman and Co., New York, NY.)

In order to express a biologically active HORP, the nucleotide sequences encoding HORP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HORP. Such elements may vary in their strength and

specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HORP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HORP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HORP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, NY, ch. 4, 8, and 16-17; and Ausubel, F.M. et al. (1995, and periodic supplements) Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HORP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus (CaMV) or tobacco mosaic virus (TMV)) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HORP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HORP can be achieved using a multifunctional E. coli vector such as Bluescript® (Stratagene) or pSport1™ plasmid (GIBCO BRL). Ligation of sequences encoding HORP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J.

Biol. Chem. 264:5503-5509.) When large quantities of HORP are needed, e.g. for the production of antibodies, vectors which direct high level expression of HORP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HORP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, supra; and Grant et al. (1987) Methods Enzymol. 153:516-54; Scorer, C. A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of HORP. Transcription of sequences encoding HORP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV. (Takamatsu, N. (1987) EMBO J. 6:307-311.) Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., Hobbs, S. or Murry, L.E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, NY; pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HORP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HORP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes.

For long term production of recombinant proteins in mammalian systems, stable expression of HORP in cell lines is preferred. For example, sequences encoding HORP can be transformed into cell lines using expression vectors which may contain viral origins of replication

and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ or *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; and Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* or *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. et al (1981) J. Mol. Biol. 150:1-14; and Murry, supra.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP) (Clontech, Palo Alto, CA), ß glucuronidase and its substrate ß-D-glucuronoside, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. et al. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HORP is inserted within a marker gene sequence, transformed cells containing sequences encoding HORP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HORP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HORP and that express HORP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of HORP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HORP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Section IV; Coligan, J. E. et al. (1997 and periodic supplements) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York, NY; and Maddox, D.E. et al. (1983) J. Exp. Med. 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HORP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HORP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Pharmacia & Upjohn (Kalamazoo, MI), Promega (Madison, WI), and U.S. Biochemical Corp. (Cleveland, OH). Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HORP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HORP may be designed to contain signal sequences which direct secretion of HORP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for

post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Manassas, VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HORP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HORP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HORP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HORP encoding sequence and the heterologous protein sequence, so that HORP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel, F. M. et al. (1995 and periodic supplements) Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, ch 10. A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HORP may be achieved in vitro using the TNT™ rabbit reticulocyte lysate or wheat germ extract systems (Promega, Madison, WI). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ^{35}S-methionine.

Fragments of HORP may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of HORP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity exists among HORP and sequences and motifs characteristic of oxidoreductases. In addition, HORP is expressed in cancerous, inflamed, immunological, fetal or proliferating, reproductive, gastrointestinal, cardiovascular, endocrine, and nervous tissues. Therefore, HORP appears to play a role in neurological, autoimmune/inflammatory, reproductive, cell proliferative, endocrine, vesicle trafficking, and smooth muscle disorders.

Therefore, in one embodiment, an antagonist of HORP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HORP. Examples of such disorders include, but are not limited to, a neurological disorder, such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; akathesia, amnesia, bipolar disorder, catatonia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, paranoid psychoses, peripheral neuropathy, postherpetic neuralgia, and Tourette's disorder; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel

syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a reproductive disorder such as disorders of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, carcinoma of the male breast, and gynecomastia; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an endocrine disorder such as disorders of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma; disorders associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism; disorders associated with hyperpituitarism including acromegaly, giantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma; disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism; disorders associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease; disorders associated with hyperparathyroidism including Conn disease (chronic hypercalemia); pancreatic disorders such as Type I or Type II diabetes mellitus and associated complications; disorders associated with the adrenals such as hyperplasia, carcinoma,

or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease; disorders associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbations of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis; and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, hypergonadal disorders associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α-reductase, and gynecomastia; a vesicle trafficking disorder such as cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, and Addison's disease; gastrointestinal disorders including ulcerative colitis, gastric and duodenal ulcers; other conditions associated with abnormal vesicle trafficking, including acquired immunodeficiency syndrome (AIDS); allergies including hay fever, asthma, and urticaria (hives); autoimmune hemolytic anemia; proliferative glomerulonephritis; inflammatory bowel disease; multiple sclerosis; myasthenia gravis; rheumatoid and osteoarthritis; scleroderma; Chediak-Higashi and Sjogren's syndromes; systemic lupus erythematosus; toxic shock syndrome; traumatic tissue damage; and viral, bacterial, fungal, helminthic, and protozoal infections; and a smooth muscle disorder, such as angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia. Smooth muscle includes, but is not limited to, that of the blood vessels, gastrointestinal tract, heart, and uterus.

In another embodiment, a vector capable of expressing HORP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HORP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HORP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HORP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HORP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HORP including, but not limited to, those listed above.

In a further embodiment, an antagonist of HORP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HORP. Such a disorder may include, but is not limited to, those neurological, autoimmune/inflammatory, reproductive, cell proliferative, endocrine, vesicle trafficking, and smooth muscle disorders described above. In one aspect, an antibody which specifically binds HORP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express HORP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HORP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HORP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HORP may be produced using methods which are generally known in the art. In particular, purified HORP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HORP. Antibodies to HORP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HORP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HORP have an amino acid sequence consisting of at least about 5 amino acids, and, more

preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HORP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HORP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HORP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton D.R. (1991) Proc. Natl. Acad. Sci. 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; and Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for HORP may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HORP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing

monoclonal antibodies reactive to two non-interfering HORP epitopes is preferred, but a competitive binding assay may also be employed. (Maddox, supra.)

In another embodiment of the invention, the polynucleotides encoding HORP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HORP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HORP. Thus, complementary molecules or fragments may be used to modulate HORP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HORP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HORP. (See, e.g., Sambrook, supra; and Ausubel, supra.)

Genes encoding HORP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HORP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HORP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage

of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HORP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HORP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nature Biotechnology 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits,

monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HORP, antibodies to HORP, and mimetics, agonists, antagonists, or inhibitors of HORP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or

solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HORP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended

purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HORP or fragments thereof, antibodies of HORP, and agonists, antagonists or inhibitors of HORP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of therapeutic to toxic effects is the therapeutic index, and it can be expressed as the ED_{50}/LD_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HORP may be used for the diagnosis of disorders characterized by expression of HORP, or in assays to monitor patients being treated with HORP or agonists, antagonists, or inhibitors of HORP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HORP include methods which utilize the antibody and a label to detect HORP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HORP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HORP expression. Normal or standard values for HORP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to HORP under conditions suitable for complex formation The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of HORP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HORP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of HORP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HORP, and to monitor regulation of HORP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HORP or closely related molecules may be used to identify nucleic acid sequences which encode HORP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding HORP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably

have at least 50% sequence identity to any of the HORP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:7-12 or from genomic sequences including promoters, enhancers, and introns of the HORP gene.

Means for producing specific hybridization probes for DNAs encoding HORP include the cloning of polynucleotide sequences encoding HORP or HORP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HORP may be used for the diagnosis of a disorder associated with expression of HORP. Examples of such a disorder include, but are not limited to, a neurological disorder, such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; akathesia, amnesia, bipolar disorder, catatonia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, paranoid psychoses, peripheral neuropathy, postherpetic neuralgia, and Tourette's disorder; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus,

emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a reproductive disorder such as disorders of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, carcinoma of the male breast, and gynecomastia; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an endocrine disorder such as disorders of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma; disorders associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism; disorders associated with hyperpituitarism including acromegaly, giantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma; disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism; disorders associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease;

disorders associated with hyperparathyroidism including Conn disease (chronic hypercalemia); pancreatic disorders such as Type I or Type II diabetes mellitus and associated complications; disorders associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease; disorders associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbations of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis; and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, hypergonadal disorders associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α-reductase, and gynecomastia; a vesicle trafficking disorder such as cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, and Addison's disease; gastrointestinal disorders including ulcerative colitis, gastric and duodenal ulcers; other conditions associated with abnormal vesicle trafficking, including acquired immunodeficiency syndrome (AIDS); allergies including hay fever, asthma, and urticaria (hives); autoimmune hemolytic anemia; proliferative glomerulonephritis; inflammatory bowel disease; multiple sclerosis; myasthenia gravis; rheumatoid and osteoarthritis; scleroderma; Chediak-Higashi and Sjogren's syndromes; systemic lupus erythematosus; toxic shock syndrome; traumatic tissue damage; and viral, bacterial, fungal, helminthic, and protozoal infections; and a smooth muscle disorder, such as angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia. Smooth muscle includes, but is not limited to, that of the blood vessels, gastrointestinal tract, heart, and uterus. The polynucleotide sequences encoding HORP may be used in Southern or Northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and ELISA assays; and in microarrays utilizing fluids or tissues from patients to detect altered HORP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HORP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HORP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization

complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HORP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HORP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HORP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HORP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HORP, or a fragment of a polynucleotide complementary to the polynucleotide encoding HORP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of HORP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and

interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; and Duplaa, C. et al. (1993) Anal. Biochem. 229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HORP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, R.A. (ed.) Molecular Biology and Biotechnology, VCH Publishers New York, NY, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding HORP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as

mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HORP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HORP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with HORP, or fragments thereof, and washed. Bound HORP is then detected by methods well known in the art. Purified HORP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HORP specifically compete with a test compound for binding HORP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HORP.

In additional embodiments, the nucleotide sequences which encode HORP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below,

in particular U.S. Ser. No. 60/091,177, filed 06/30/98, and U.S. Ser. No.[Atty Docket # PF-0557 P], filed 07/16/98, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

HMC1NOT01

The human mast cell HMCINOT01 cDNA library was custom constructed by Stratagene (Stratagene, La Jolla, CA) using mRNA purified from cultured HMC-1 cells. The human mast cell (HMC-1) cDNA library was prepared by purifying poly(A+)RNA (mRNA) from human mast cells and then enzymatically synthesizing double stranded complementary DNA (cDNA) copies of the mRNA. Synthetic adaptor oligonucleotides were ligated onto the ends of the cDNA enabling its insertion into the lambda vector. The HMC-1 library was constructed using the UNIZAP vector system (Stratagene).

The HMC-1 cDNA library was screened with either DNA probes or antibody probes, and the PBLUESCRIPT phagemid (Stratagene) was excised in vivo. The custom-constructed library phage particles were infected into E. coli host strain XL1-BLUE cells (Stratagene).

PROSNON01

The PROSNON01 normalized cDNA library was constructed from microscopically normal prostate tissue obtained from a 28-year-old Caucasian male who died from a gun shot wound.

The frozen tissue was homogenized and lysed using a Brinkmann Homogenizer Polytron PT-3000 (Brinkmann Instruments, Westbury, NY) in guanidinium isothiocyanate solution. The lysate was centrifuged over a 5.7 M CsCl cushion using an Beckman SW28 rotor in a Beckman L8-70M Ultracentrifuge (Beckman Instruments) for 18 hours at 25,000 rpm at ambient temperature. The RNA was extracted with acid phenol pH 4.7, precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in RNAse-free water, and treated with DNase at 37°C. The RNA extraction and precipitation were repeated as before. The mRNA was then isolated using the OLIGOTEX kit (QIAGEN, Valencia, CA) and used to construct the cDNA library.

The RNA was handled according to the recommended protocols in the SUPERSCRIPT plasmid system (Life Technologies). The cDNAs were fractionated on a Sepharose CL4B column (Amersham Pharmacia Biotech), and those cDNAs exceeding 400 bp were ligated into PSPORT 1 (Life Technologies). The plasmid PSPORT1 was subsequently transformed into DH5α™

competent cells (Life Technologies).

The construction of the cDNA libraries from which SEQ ID NO:7-10 were isolated was as follows. RNA was purchased from Clontech (Palo Alto, CA) or isolated at Incyte from tissues described in Table 4. The tissue was homogenized and lysed in guanidinium isothiocyanate, and the lysate was centrifuged over a CsCl cushion. Alternatively, the tissue was homogenized and lysed in phenol or a suitable mixture of denaturants such as TRIZOL reagent (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate, and the lysate was extracted with chloroform (1:5 v/v). RNA was precipitated from lysates with either isopropanol or sodium acetate and ethanol. Alternatively, RNA was purified from lysates by preparative agarose gel electrophoresis and recovered from Whatman P81 paper (Whatman, Lexington, MA). Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity, and RNA was maintained in RNase-free solutions. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega, Madison, WI), Oligotex resin, or the Oligotex kit (QIAGEN Inc, Chatsworth, CA). Alternatively, RNA was isolated directly from tissue lysates using the RNA Isolation kit (Stratagene) or the Ambion PolyA Quick kit (Ambion, Austin, TX).

RNA was used for cDNA synthesis and construction of the cDNA libraries according to procedures recommended in the UNIZAP™ vector (Stratagene, La Jolla, CA) or SuperScript plasmid system (Life Technologies, Inc), both of which are based on methods well known in the art (Ausubel, 1997, units 5.1-6.6). Alternatively, cDNA libraries were constructed by Stratagene using RNA provided by Incyte. Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and cDNA was digested with an appropriate restriction enzyme(s). For most libraries, cDNA was size-selected (300-1000 bp) using Sephacryl S1000 or Sepharose CL2B or CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT (Stratagene), PSPORT 1 (Life Technologies), pINCY (Incyte Pharmaceuticals Inc, Palo Alto, CA). pINCY was amplified in JM109 cells and purified using the QiaQuick column (QIAGEN Inc). Recombinant plasmids were transformed into competent E. coli cells, e.g., XL1-Blue, XL1-BlueMRF, or SOLR (Stratagene) or DH5α, DH10B, or ElectroMAX DH10B (Life Technologies).

II. Isolation of cDNA Clones

HMC1NOT01

The phagemid forms of individual cDNA clones were obtained by the in vivo excision

process, in which the host bacterial strain was coinfected with both the lambda library phage and an f1 helper phage (UNIZAP vector system, Stratagene). Proteins derived from both the library-containing phage and the helper phage nicked the lambda DNA, initiated new DNA synthesis from defined sequences on the lambda target DNA and created a smaller, single stranded circular phagemid DNA molecule that included all DNA sequences of the PBLUESCRIPT (Stratagene) plasmid and the cDNA insert. The phagemid DNA was secreted from the cells and purified, then used to re-infect fresh host cells, where the double stranded phagemid DNA was produced. Because the phagemid carried the gene for ß-lactamase, the newly-transformed bacteria were selected on medium containing ampicillin.

Phagemid DNA was purified using the MAGIC MINIPREPS DNA purification system (Promega, Madison, WI). The DNA was eluted from the purification resin already prepared for DNA sequencing and other analytical manipulations. Phagemid DNA may also be purified using the QIAWELL-8 Plasmid or DNA Purification system (QIAGEN). The DNA was eluted from the purification resin and prepared for DNA sequencing and other analytical manipulations.

An alternative method for purifying phagemid DNA utilizes the Miniprep Kit available from Advanced Genetic Technologies Corp. (Gaithersburg, MD). This kit is in the 96-well format and provides enough reagents for 960 purifications. Each kit contains a recommended protocol, which is employed except for the following changes. First, the 96 wells are each filled with only 1 ml of sterile Terrific Broth (Life Technologies) with carbenicillin at 25 mg/L and glycerol at 0.4%. After the wells are inoculated, the bacteria are cultured for 24 hours and lysed with 60 μl of lysis buffer. A centrifugation step (2900 rpm for 5 minutes) is performed before the contents of the block is added to the primary filter plate. The optional step of adding isopropanol to TRIS buffer is not performed. After the last step in the protocol, samples are transferred to a Beckman 96-well block for storage.

PROSNON01

4.4×10^6 independent clones of a normal prostate plasmid library in E. coli strain DH12S (Life Technologies) were grown in liquid culture under carbenicillin (25mg/L) and methicillin (1mg/ml) selection following transformation by electroporation. The culture was checked spectrophotometrically (Model DU-7 Spectrophotometer, Beckman Instruments) and allowed to grow to an OD600 of 0.2 and then superinfected with a 5-fold excess of the helper phage M13K07 according to the method of Vieira et al. (1987, Methods Enzymol. 153:3-11), herein incorporated by reference.

To reduce the number of excess cDNA copies according to their abundance levels in the library, the cDNA library was then normalized in a single round according to the procedure of

Soares et al. (1994 Proc. Natl. Acad. Sci. 91:9928-9932), herein incorporated by reference, with the following modifications. The primer to template ratio in the primer extension reaction was increased from 2:1 to 10:1. The ddNTP concentration in this reaction was reduced to 150μM each ddNTP to allow the generation of longer (400-1000nt) primer extension products. The reannealing hybridization was extended from 13 to 19 hours. The single stranded DNA circles of the normalized library were purified by hydroxyapatite chromatography, converted to partially double-stranded by random priming, and electroporated into DH10B competent bacteria (Life Technologies).

Plasmid DNA was released from the cells and purified using the R.E.A.L. PREP 96 plasmid kit (QIAGEN, Valencia, CA). This kit enabled the simultaneous purification of 96 samples in a 96-well block using multi-channel reagent dispensers. The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Life Technologies) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) after inoculation, the cultures were incubated for 19 hours and at the end of incubation, the cells were lysed with 0.3 ml of lysis buffer; and 3) following isopropanol precipitation, the plasmid DNA pellet was resuspended in 0.1 ml of distilled water. After the last step in the protocol, samples were transferred to a 96-well block for storage at 4° C.

Isolation of cDNA clones containing SEQ ID NO:7-10 was as follows. Plasmids were recovered from host cells by in vivo excision (UniZAP vector system, Stratagene) or by cell lysis. Plasmids were purified using the MAGIC MINIPREPS DNA purification system (Promega, Madison, WI); Miniprep kit (Advanced Genetic Technologies Corporation, Gaithersburg, MD); QIAwell-8 Plasmid, QIAwell PLUS DNA, or QIAwell ULTRA DNA purification systems; or R.E.A.L. PREP 96 plasmid kit (QIAGEN Inc) using the recommended protocol. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR (Rao, V.B. (1994) Anal. Biochem. 216:1-14) in a high-throughput format. Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates (Genetix Ltd, Christchurch UK) and concentration of amplified plasmid DNA was quantified fluorometrically using Pico Green Dye (Molecular Probes, Eugene OR) and a Fluoroscan II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

The cDNAs were prepared for sequencing using either an ABI PRISM™ CATALYST™

800 (Perkin-Elmer Applied Biosystems, Foster City, CA) or a MICROLAB® 2200 (Hamilton Co., Reno, NV) sequencing preparation system in combination with Peltier PTC-200 thermal cyclers (MJ Research, Inc., Watertown, MA). The cDNAs were sequenced using the ABI PRISM™ 373 or 377 sequencing systems and ABI protocols, base calling software, and kits (Perkin-Elmer Applied Biosystems). Alternatively, solutions and dyes from Amersham Pharmacia Biotech, Ltd. were used in place of the ABI kits. In some cases, reading frames were determined using standard methods (Ausubel, supra). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA, extension, and shotgun sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the software programs used, corresponding algorithms, references, and cutoff parameters used where applicable. The references cited in the third column of Table 5 are incorporated by reference herein. Sequence alignments were also analyzed and produced using MACDNASIS PRO software (Hitachi Software Engineering Co., Ltd. San Bruno, CA) and the multisequence alignment program of LASERGENE software (DNASTAR Inc, Madison WI).

The polynucleotide sequences were validated by removing vector, linker, and polyA tail sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation, using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. This was followed by translation of the full length polynucleotide sequences to derive the corresponding full length amino acid sequences. These full length polynucleotide and amino acid sequences were subsequently analyzed by querying against databases such as the GenBank databases described above and SwissProt, BLOCKS, PRINTS, PFAM, and Prosite.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; and Ausubel, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST are used to search for identical or

related molecules in nucleotide databases such as GenBank or LIFESEQ™ database (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar.

The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of Northern analysis are reported as a list of libraries in which the transcript encoding HORP occurs. Abundance and percent abundance are also reported. Abundance directly reflects the number of times a particular transcript is represented in a cDNA library, and percent abundance is abundance divided by the total number of sequences examined in the cDNA library.

V. Extension of HORP Encoding Polynucleotides

The nucleic acid sequences of Incyte Clones 321510, 634343, 1942326, 2395269, 008879, and 2274011 were used to design oligonucleotide primers for extending partial nucleotide sequences to full length. For each nucleic acid sequence, one primer was synthesized to initiate extension of an antisense polynucleotide, and the other was synthesized to initiate extension of a sense polynucleotide. Primers were used to facilitate the extension of the known sequence "outward" generating amplicons containing new unknown nucleotide sequence for the region of interest. The initial primers were designed from the cDNA using OLIGO™ 4.06 (National Biosciences, Plymouth, MN), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries (GIBCO BRL) were used to extend the sequence. If more than one extension is necessary or desired, additional sets of primers are designed to further extend the known region.

High fidelity amplification was obtained by following the instructions for the XL-PCR™ kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix. PCR was performed

using the Peltier Thermal Cycler (PTC200; M.J. Research, Watertown, MA), beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, with the following parameters:

Step 1	94° C for 1 min (initial denaturation)
Step 2	65° C for 1 min
Step 3	68° C for 6 min
Step 4	94° C for 15 sec
Step 5	65° C for 1 min
Step 6	68° C for 7 min
Step 7	Repeat steps 4 through 6 for an additional 15 cycles
Step 8	94° C for 15 sec
Step 9	65° C for 1 min
Step 10	68° C for 7:15 min
Step 11	Repeat steps 8 through 10 for an additional 12 cycles
Step 12	72° C for 8 min
Step 13	4° C (and holding)

A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a low concentration (about 0.6% to 0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products were excised from the gel, purified using QIAQUICK™ (QIAGEN Inc.), and trimmed of overhangs using Klenow enzyme to facilitate religation and cloning.

After ethanol precipitation, the products were redissolved in 13 μl of ligation buffer, 1μl T4-DNA ligase (15 units) and 1μl T4 polynucleotide kinase were added, and the mixture was incubated at room temperature for 2 to 3 hours, or overnight at 16° C. Competent E. coli cells (in 40 μl of appropriate media) were transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium. (See, e.g., Sambrook, supra, Appendix A, p. 2.) After incubation for one hour at 37°C, the E. coli mixture was plated on Luria Bertani (LB) agar (See, e.g., Sambrook, supra, Appendix A, p. 1) containing carbenicillin (2x carb). The following day, several colonies were randomly picked from each plate and cultured in 150 μl of liquid LB/2x carb medium placed in an individual well of an appropriate commercially-available sterile 96-well microtiter plate. The following day, 5 μl of each overnight culture was transferred into a non-sterile 96-well plate and, after dilution 1:10 with water, 5 μl from each sample was transferred into a PCR array.

For PCR amplification, 18 μl of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for

the extension reaction were added to each well. Amplification was performed using the following conditions:

Step 1	94° C for 60 sec
Step 2	94° C for 20 sec
Step 3	55° C for 30 sec
Step 4	72° C for 90 sec
Step 5	Repeat steps 2 through 4 for an additional 29 cycles
Step 6	72° C for 180 sec
Step 7	4° C (and holding)

Aliquots of the PCR reactions were run on agarose gels together with molecular weight markers. The sizes of the PCR products were compared to the original partial cDNAs, and appropriate clones were selected, ligated into plasmid, and sequenced.

In like manner, the nucleotide sequences of SEQ ID NO:7-12 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for 5' extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:7-12 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO™ 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham, Chicago, IL), and T4 polynucleotide kinase (DuPont NEN®, Boston, MA). The labeled oligonucleotides are substantially purified using a Sephadex™ G-25 superfine size exclusion dextran bead column (Pharmacia & Upjohn, Kalamazoo, MI). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba 1, or Pvu II (DuPont NEN, Boston, MA).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham, NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR™ film (Kodak, Rochester, NY) is exposed to the blots for several hours, hybridization patterns are compared visually.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE™. Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; and Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the HORP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HORP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO™ 4.06 software and the coding sequence of HORP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HORP-encoding transcript.

IX. Expression of HORP

Expression and purification of HORP is achieved using bacterial or virus-based expression systems. For expression of HORP in bacteria, cDNA is subcloned into an appropriate

vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HORP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HORP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HORP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E. K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HORP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Pharmacia, Piscataway, NJ). Following purification, the GST moiety can be proteolytically cleaved from HORP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak, Rochester, NY). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN Inc, Chatsworth, CA). Methods for protein expression and purification are discussed in Ausubel, F. M. et al. (1995 and periodic supplements) Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, ch 10, 16. Purified HORP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of HORP Activity

The assay for human oxidoreductase proteins is based upon the increase in extinction coefficient of NAD(P)H coenzyme at 340 nm for the measurement of oxidation activity, or the decrease in extinction coefficient of NAD(P)H coenzyme at 340 nm for the measurement of reduction activity (Dalziel, K. (1963) J. Biol. Chem. 238:2850-2858). One of three substrates may

be used: Asn-βGal, biocytidine, or ubiquinone-10. The respective subunits of the enzyme reaction, for example, cytochtome c_1-b oxidoreductase and cytochrome c, are reconstituted. The reaction mixture contains 1-2 mg/ml HORP, 15 mM substrate, 2.4 mM $NAD(P)^+$ or 2.0 mM NAD(P)H, in 0.1 M phosphate buffer, pH 7.1 ($NAD(P)^+$, oxidation reaction), or 0.1 M Na_2HPO_4 buffer, pH 7.4 (NAD(P)H, reduction reaction), in a total volume of 0.1 ml. Changes in absorbance at 340 nm (A_{340}) are measured at 23.5° C using a recording spectrophotometer (Shimadzu Scientific Instruments, Inc., Pleasanton, CA). The amount of NAD(P)H is stoichiometrically equivalent to the amount of substrate initially present, and the change in A_{340} is a direct measure of the amount of NAD(P)H produced; $\Delta A_{340} = 6620[NADH]$. HORP activity is proportional to the amount of NAD(P)H present in the assay.

Alternatively, HORP-5 activity is demonstrated by the ability to convert linoleic acid to cis-linolenic acid in purified microsomes. (Ivanetich, K.M. et al. (1996) Biochim. Biophys. Acta 1292:120-132.). Endogenous microsomal acyl-Coenzyme A (CoA) synthetase converts linoleic acid to linoleoyl-CoA, the direct substrate for HORP-5. HORP-5 converts linoleoyl-CoA to linolenoyl-CoA, which is subsequently converted to 2-γ-linolenoyl-phospholipid by endogenous microsomal lysophospholipid acyltransferase. Saponification converts 2-γ-linolenoyl-phospholipid to cis-linolenic acid, which is measured as "product" in this assay. Microsomes, purified from mammalian cells transfected with HORP-5 cDNA and from control cells, are suspended in Buffer A (0.05 M potassium phosphate, pH 7.4 containing 0.25 M sucrose, 0.15 M KCl, 1.5 mM glutathione, 5 mM $MgCl_2$, and 4 mM EDTA. The microsomes (0.5 mg/ml protein) are incubated in Buffer A containing 7.5 mM ATP, 1 mM Coenzyme A, 2.6 mM NADH, 40 mM KF, 0.33 mM nicotinamide, bovine serum albumin (11.5 μg/μg added fatty acid), and [1-^{14}C] linoleic acid (0.45-10.9 nmol, 26-632 nCi) in 1 ml total volume. Incubations are at 35°C with shaking at 60 cycles/min for times varying from 0 to 10 min. The reactions are terminated with the addition of 1 ml 10% potassium hydroxide in methanol containing 0.005% butylated hydroxytoluene.

Reaction mixtures, containing 1 mg of fatty acid carriers, are saponified for 30 min at 60°C under argon, acidified and extracted three times with 2 ml hexane. The hexane extract is dried under nitrogen at approximately 45°C, fatty acids are dissolved in 0.5 ml ethanol and the solution passed through a 0.45 μm filter. The fatty acid substrate and product are separated on a high performance liquid chromatography column (Zorbax ODS, DuPont, Wilmington, DE, 25 cm by 0.45 cm; or Spherisorb ODS, Phase Separations, UK, 25 cm by 0.45 cm, 10 μm pore size). The columns are run in acetonitrile:30 mM phosphoric acid (65:35, v/v) at a flow rate of 2 ml/minute. Fractions of 2 ml are collected and the radioactivity in each fraction determined by liquid scintillation counting. The retention times on the Zorbax ODS (run at 35°C) and Spherisorb

(run at 25°C) are for linoleic acid (substrate), 24 +/- 2 and 15+/-1 minutes, respectively, and for cis-linolenic acid (product) 16+/-1 and 11+/-1 minutes, respectively. Delta-6 desaturase activity is calculated from the ratio: dpm product/(dpm product + dpm substrate). The activity present in the microsomes from HORP-5 transfected cells over that present in control cells is due to the HORP-5.

In a further alternative, HORP-6 activity is demonstrated by the ability to catalyze transmembrane electron transfer in a reconstituted phospholipid vesicle membrane. (Srivastava, M. et al. (1984) J. Biol. Chem. 259:8072-8075.) The assay measures the transfer of electrons from ascorbate inside the vesicles to ferricytochrome c in the external medium. To prepare the vesicles, egg phosphatidylcholine (4 mg, Sigma) in hexane is dried under nitrogen. To the dried lipid, 100 µl of 10% cholate is added and sonicated in a bath sonicator until the solution is clear. HORP-6 in 1% octylglucoside (Sigma) is mixed rapidly with the sonicated mixture. The ascorbate-loaded vesicles are prepared by dialyzing the HUCY-2/lipid/detergent mixture against 200 ml of 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (Hepes) pH 7.4, 0.05 M NaCl, 0.1 M ascorbic acid under nitrogen, at 4°C for 24 hours. To accelerate the removal of cholate and octylglucoside, 4 g of SM-2 Bio-Beads (Bio-Rad) are packed in a separate dialysis bag and added to the Hepes/salt buffer solution. Normally 1-2 nmol of HORP-6 are used and the molar ratio of octylglucoside to cholate to phosphatidylcholine is approximately 7:6:1. The HORP-6 to phospholipid ratio is approximately 1:5000. Separation of external ascorbate from the vesicles is accomplished by centrifuging the sample (typically 0.6 ml) through hydrated Sephadex G-25-80 (5 ml bed volume) and collecting the purified vesicles that go through the Sephadex column. (Fry, D.W. et al. (1978) Anal. Biochem. 90:809-815.)

Ferricytochrome c is prepared by adding potassium ferricyanide (4 mM) to 2 mM horse cytochrome c (Sigma). Ferricyanide is removed by dialyzing the solution against 200 volumes of 0.3 M sucrose, 10 mM Hepes/KOH, pH 7.0 at 4°C for 24 hours and then passing through a Sephadex G-25 column equilibrated with 50 mM Hepes, pH 7.4, 0.15 M NaCl. The ferricytochrome c is added to the vesicles. Electron transfer from ascorbate trapped inside the vesicles reduces the ferricytochrome c, causing the absorbance at 550 nm of the vesicle solution to increase. This increase in absorbance at 550 nm, measured using a spectrophotometer, is proportional to the amount of HORP-6 in the vesicles.

XI. Functional Assays

HORP function is assessed by expressing the sequences encoding HORP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT™ (Life Technologies, Gaithersburg, MD)

and pCR™ 3.1 (Invitrogen, Carlsbad, CA, both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP) (Clontech, Palo Alto, CA), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP, and to evaluate properties, for example, their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York, NY.

The influence of HORP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HORP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success, NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HORP and other genes of interest can be analyzed by Northern analysis or microarray techniques.

XII. Production of HORP Specific Antibodies

HORP substantially purified using polyacrylamide gel electrophoresis (PAGE)(see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the HORP amino acid sequence is analyzed using LASERGENE™ software (DNASTAR Inc.) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in

hydrophilic regions are well described in the art. (See, e.g., Ausubel supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry and coupled to KLH (Sigma, St. Louis, MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring HORP Using Specific Antibodies

Naturally occurring or recombinant HORP is substantially purified by immunoaffinity chromatography using antibodies specific for HORP. An immunoaffinity column is constructed by covalently coupling anti-HORP antibody to an activated chromatographic resin, such as CNBr-activated Sepharose (Pharmacia & Upjohn). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HORP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HORP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HORP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HORP is collected.

XIV. Identification of Molecules Which Interact with HORP

HORP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton et al. (1973) Biochem. J. 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HORP, washed, and any wells with labeled HORP complex are assayed. Data obtained using different concentrations of HORP are used to calculate values for the number, affinity, and association of HORP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	7	321510	EOSIHET02	321510H1 (EOSIHET02), 2593359H1 (LUNGNOT22), 2902337T6 (DRGCNOT01), 3016872H1 (MUSCNOT07), 3513795F6 (ENDINOT02)
2	8	634343	NEUTGMT01	562277R6 (NEUTLPT01), 634343H1 (NEUTGMT01), 634343X12 (NEUTGMT01), 634343X14 (NEUTGMT01), 634343X17 (NEUTGMT01), 1217756T1 (NEUTGMT01)
3	9	1942326	HIPONOT01	598737R1 (BRSTNOT02), 891269T6 (STOMTUT01), 926750T1 (BRAINOT04), 1473641T1 (LUNGTUT03), 1942326H1 (HIPONOT01), 1942326T6 (HIPONOT01), 3056158H1 (LNODNOT08)
4	10	2395269	THP1AZT01	1281240F6 (COLNNOT16), 1281240T6 (COLNNOT16), 1508803F6 (LUNGNOT14), 1709112F6 (PROSNOT16), 2395269H1 (THP1AZT01)
5	11	008879	HMC1NOT01	008879H1 (HMC1NOT01), 784197R1 (PROSNOT05), 869194R1 (LUNGAST01), 1319527F1 (BLADNOT04), 1320867F1 (BLADNOT04), 1457227F1 (COLNFET02), 1593214X19C1 (BRAINOT14), 1593214X20C1 (BRAINOT14), 1593214X21C1 (BRAINOT14), 1986117R6 (LUNGAST01), 3491654H1 (ADRETUT07)
6	12	2274011	PROSNON01	056483R1 (FIBRNOT01), 081829F1 (SYNORAB01), 099898R6 (ADRENOT01), 1250153F1 (LUNGFET03), 1251167F1 (LUNGFET03), 1282535F1 (COLNNOT16), 1307709F6 (COLNFET02), 1362055F1 (LUNGNOT12), 1362485T6 (LUNGNOT12), 1442809F1 (THYRNOT03), 1511496F1 (LUNGNOT14), 2274011H1 (PROSNON01), SAUA01554F1

Table 2

Protein SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification	Analytical Methods
1	310	S15 S58 T91 T98 S111 S145 S193 S216 S108 S136	N252		N-terminal asparagine amidohydrolase	BLAST
2	520	S46 T78 T210 T246 S302 T424 T326 T336 S370 S423	N39 N273 N347 N357 N411 N468	T41 to K87 S184 to T236	biotinidase; biotin-amide amidohydrolase	BLAST, HMM
3	349	S9 S10 T11 T118 S125 S178 S208 S230 T260	N291	L59 to V86 K146 to I192	quinone oxidoreductase	BLAST, BLOCKS
4	332	S260 S24 T100 S220 S174 T201 Y273		G73 to G92 G242 to 259	Pig3	BLAST, BLOCKS, PRINTS, PFAM, HMM

Table 2 (cont)

Protein SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification	Analytical Methods
5	444	S22 T119 S429 S433	N45	Transmembrane domains: F133 to V151 F149 to A169 V300 to L323 Heme-binding domain: T20 to E98 Cytochrome b5 signature: V43 to H53 H53 to D67 A68 to F75 I109 to L116 F322 to H332 Q27 to H76	cytochrome b5/desaturase fusion protein	BLAST, BLOCKS, PRINTS, PFAM, HMM
6	286	S43 S185 S260 T74 S150 T179 S185	N231	Signal sequence: M1 to A27 Transmembrane domains: W7 to F26 V52 to P72 S84 to F105 V160 to T179 V201 to T220	cytochrome b561	BLAST, HMM, SPSCAN

Table 3

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease Class (Fraction of Total)	Vector
7	Reproductive (0.205) Gastrointestinal (0.182) Cardiovascular (0.136)	Cancer (0.432) Inflammation (0.364) Fetal (0.136)	PBLUESCRIPT
8	Hematopoietic/Immune (0.421) Gastrointestinal (0.316) Reproductive (0.158)	Inflammation (0.474) Cancer (0.421) Other (0.053)	PSPORT 1
9	Reproductive (0.256) Cardiovascular (0.154) Gastrointestinal (0.128)	Cancer (0.487) Inflammation (0.205) Fetal (0.179)	PBLUESCRIPT
10	Reproductive (0.265) Cardiovascular (0.204) Gastrointestinal (0.163)	Cancer (0.531) Fetal (0.184) Inflammation (0.163)	pINCY
11		Cancer (0.590) Inflammation (0.260) Fetal (0.230)	PBLUESCRIPT
12		Cancer (0.480) Inflammation (0.320) Fetal (0.120)	PSPORT 1

Table 4

Nucleotide SEQ ID NO:	Clone ID	Library	Library Comment
7	321510	EOSIHET02	The EOSIHET02 library was constructed using RNA isolated from peripheral blood cells apheresed from a 48-year-old Caucasian male. Patient history included hypereosinophilia. The cell population was determined to be greater than 77% eosinophils by Wright's staining.
8	634343	NEUTGMT01	The NEUTGMT02 library was constructed using RNA isolated from peripheral blood granulocytes isolated from buffy coat units obtained from 20 unrelated male and female donors. Cells were cultured in 10 nM GM-CSF for 1 hour before washing and harvesting for total RNA preparation.
9	1942326	HIPONOT01	The HIPONOT01 library was constructed using RNA isolated from the hippocampus tissue of a 72-year-old Caucasian female who died from an intracranial bleed. Patient history included nose cancer, hypertension, and arthritis.
10	2395269	THP1AZT01	The THP1AZT01 library was constructed using RNA isolated from THP-1 promonocyte cells treated for three days with 0.8 micromolar 5-aza-2'-deoxycytidine. THP-1 (ATCC TIB 202) is a human promonocyte line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (ref: Int. J. Cancer (1980) 26:171).

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS and PRINTS databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and Probability value= 1.0E-3 or less
PFAM	A Hidden Markov Models-based application useful for protein family search.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits, depending on individual protein families

Table 5 (cont)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Score=4.0 or greater
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	